Improved annual rainfall-runoff forecasting using PSO–SVM model based on EEMD

Wen-chuan Wang, Dong-mei Xu, Kwok-wing Chau and Shouyu Chen

ABSTRACT

Rainfall-runoff simulation and prediction in watersheds is one of the most important tasks in water resources management. In this research, an adaptive data analysis methodology, ensemble empirical mode decomposition (EEMD), is presented for decomposing annual rainfall series in a rainfall-runoff model based on a support vector machine (SVM). In addition, the particle swarm optimization (PSO) is used to determine free parameters of SVM. The study data from a large size catchment of the Yellow River in China are used to illustrate the performance of the proposed model. In order to measure the forecasting capability of the model, an ordinary least-squares (OLS) regression and a typical three-layer feed-forward artificial neural network (ANN) are employed as the benchmark model. The performance of the models was tested using the root mean squared error (RMSE), the average absolute relative error (AARE), the coefficient of correlation (*R*) and Nash–Sutcliffe efficiency (NSE). The PSO–SVM–EEMD model improved ANN model forecasting (65.99%) and OLS regression (64.40%), and reduced RMSE (67.7%) and AARE (65.38%) values. Improvements of the forecasting results regarding the *R* and NSE are 8.43%, 18.89% and 182.7%, 164.2%, respectively. Consequently, the presented methodology in this research can enhance significantly rainfall-runoff forecasting at the studied station.

Key words | annual rainfall-runoff, artificial neural networks, ensemble empirical mode decomposition, forecasting, particle swarm optimization, support vector machine

Wen-chuan Wang (corresponding author)
Dong-mei Xu
Department of Water Conservancy,
North China University of Water Resources and Electric Power,
Zhengzhou 450011,
China
E-mail: *wangwen1621@163.com*

Dong-mei Xu
Shouyu Chen
Water Resources and Flood Control,
Dalian University of Technology,
Dalian 116085,
China

Kwok-wing Chau
Department of Civil and Environmental Engineering,
Hong Kong Polytechnic University,
Hung Hom,
Kowloon,
Hong Kong,
China

INTRODUCTION

Runoff simulation and prediction in watersheds is a prerequisite for many practical applications involving conservation, environmental disposal and water resources management. Yet, the rainfall-runoff process is a complex, non-linear and dynamic hydrological phenomenon to simulate due to the spatial–temporal variability and inter-relationships of underlying climatic and physiographic variables (Zhang & Govindaraju 2003). The development of rainfall-runoff models has undergone substantial changes since Sherman pioneered the unit hydrograph theory in 1932 (Liong *et al.* 2002). Based on the description of the governing processes, rainfall-runoff models can be classified as either physically based (knowledge-driven) or system theoretic (data-driven). Physically based models involve a detailed interaction of various physical processes controlling the hydrologic behavior of a system. However, system theoretic models are instead based primarily on observations (measured data) and seek to characterize the system response from those data using transfer functions (Wu & Chau 2011). In recent years, data-driven modeling approaches are being widely used as surrogate for physically based models, as they overcome some limitations associated with physically based approaches.

As an example of data-driven models, artificial neural network (ANN) techniques are advocated as an appropriate and sensible method for the combination of simulated river flows of a suite of rainfall-runoff models (Shamseldin *et al.* 1997), as a successful tool, to solve various problems concerned with hydrology and water resources engineering (ASCE 2000a, b) mainly owing to their capability to treat

doi: 10.2166/hydro.2013.134

complicated and non-linear problems and reproduce the highly non-linear nature of relationship between hydrological variables. ANN use in modeling the rainfall-runoff process started with a preliminary study by Halff *et al.* (1993). Since then, many studies in the context of rainfall-runoff modeling using ANNs have been employed as alternative tools in developing data-driven models of hydrological processes.

Extensive reviews of previous works on ANN applications in hydrologic simulation and forecasting have been reported in ASCE (2000a, b), Maier & Dandy (2000) and Dawson & Wilby (2001). These studies have demonstrated the capability of ANNs in runoff forecasting and have proved that ANNs may offer a promising alternative for rainfall-runoff modeling (Karunanithi *et al.* 1994; Hsu *et al.* 1995; Minns & Hall 1996; Shamseldin 1997; Braddock *et al.* 1998; Dawson & Wilby 1998; Sajikumar & Thandaveswara 1999; Tokar & Johnson 1999). More recently, Anmala *et al.* (2000) used feed forward ANN for monthly rainfall-runoff simulations and obtained encouraging results for the examined basins. Zhang & Govindaraju (2000) performed monthly runoff forecasting from monthly precipitation and temperature, for three medium-sized basins in Kansas, USA, using a feed-forward ANN. Abrahart (2003) used a continuous single model bootstrap to neural network for rainfall-runoff forecasting. Sarangi *et al.* (2005) associated selected geomorphological parameters as inputs with rainfall duration for prediction of surface runoff. Jeong & Kim (2005) compared single neural network (SNN), ensemble neural network (ENN) and conceptual rainfall-runoff model for ensemble stream flow predictions and found ENN to perform the best among the three models. Meanwhile, various training methods available were investigated and compared in training ANNs for modeling the rainfall-runoff process (Srinivasulu & Jain 2006; Sedki *et al.* 2009; Lin & Wu 2011). These methods included the popular back-propagation algorithm (BPA), genetic algorithm (RGA), self-organizing map (SOM), and a two-step learning algorithm. Two hybrid artificial intelligence (AI) based models are introduced by Nourani *et al.* (2011) for watershed rainfall-runoff modeling. Lohani *et al.* (2011) compared ANN, fuzzy logic (FL) and linear transfer function (LTF)-based approaches for daily rainfall-runoff modeling. The wavelet and ANN techniques were employed to form a loose type of wavelet ANN hybrid model (NW) for 1-, 2- and 3-day flow forecasting at a site of the Brahmani River, India (Pramanik *et al.* 2011).

Support vector machines (SVMs), which are based on the statistical learning theory and were introduced by Vapnik (1995), are a relatively new class of models in the data-driven prediction field. Although SVMs have remarkable successes in various fields, there are few studies on their applications in water resources and hydrology (Behzad *et al.* 2009). Liong & Sivapragasam (2002) demonstrated that SVM models show good generalization performance in their applications on flood forecasting and rainfall-runoff modeling. Bray & Han (2004) described an exploration in using SVM models in flood forecasting. A SVM model was used to forecast flows at different time scales: seasonal flow volumes, hourly stream flows and long-term discharges (Asefa *et al.* 2006; Lin *et al.* 2006). Wu *et al.* (2008) applied it successfully to water level prediction. The autoregressive moving-average (ARMA) models, ANNs approaches, adaptive neural-based fuzzy inference system (ANFIS) techniques, genetic programming (GP) models and SVM method were examined by Wang *et al.* (2009) for forecasting monthly discharge time series. Yoon *et al.* (2011) developed two non-linear time-series models for predicting groundwater level fluctuations using ANNs and SVMs. An improved SVM model with adaptive insensitive factor was proposed by Guo *et al.* (2011) in predicting monthly stream flow. Samsudin *et al.* (2011) proposed a novel hybrid forecasting model known as GLSSVM, which combines the group method of data handling (GMDH) and the least squares support vector machine (LSSVM) for river flow forecasting.

The empirical mode decomposition (EMD), as a new data-preprocessing technique, was first introduced by Huang *et al.* (1998). It is an empirical, intuitive, direct and self-adaptive data processing method, which was proposed especially for non-linear and non-stationary data. EMD has been proven to be a quite effective method for extracting signals from data generated in noisy non-linear and non-stationary processes and was successfully applied in many areas, such as solar cycle, earthquake engineering, crude oil price analysis, biomedical signals, speaker identification system, and so on (Coughlin & Tung 2004; Spanos *et al.* 2007; Zhang *et al.* 2008; Wu & Huang 2009;

Karagiannis & Constantinou 2011; Wu & Tsai 2011). The main advantage of EMD over traditional approaches is its complete self-adaptiveness and its very local ability both in physical space and frequency space (Huang *et al.* 2009). The hydrology data (rainfall processes, runoff processes) have non-linear and non-stationary characteristics, but direct applications of the EMD method in analysis to non-linear hydrology data are very few (Huang *et al.* 2009). Motivated by the idea of 'decomposition and ensemble' (Yu *et al.* 2008; Guo *et al.* 2012), the original time series can be decomposed into several sub-series. Each sub-series can be forecasted with the purpose of easy predication tasks and fine results, and the final forecasted value can be obtained by summing the forecasted value of each sub-series. Hence, the ensemble empirical mode decomposition (EEMD) (Wu & Huang 2009) is applied to deal with rainfall data, with a complete decomposition resulting in four intrinsic models from high to low frequency and trend for aiding the prediction of runoff. The idea is inspired by the hybrid method which can overcome the drawbacks of individual models and to generate a synergetic effect in forecasting. In this paper, the particle swarm optimization (PSO)–SVM model based on the EEMD model is presented for rainfall-runoff modeling. The EEMD is used for decomposing annual rainfall series in a rainfall-runoff model based SVM and the PSO is employed to determine the free parameters of the SVMs. In addition, the development and performances of other models are also demonstrated with original annual rainfall series and decomposed annual rainfall series using EEMD before discussing the results and making concluding remarks.

The rest of the paper is organized as follows: the next section introduces the study area and data materials. This is followed by a brief introduction to the basic theory and algorithm of EEMD, SVM, PSO, ANNs and ordinary least-squares (OLS) regression. The model construction is then presented and the derived intrinsic modes and detailed analyses based on a composition of intrinsic modes are shown in this section. Optimizing the SVM parameters with PSO is also demonstrated in this section. After this, the performance evaluation of model is introduced, followed by the application results, including comparison and analyses, and finally the conclusions.

STUDY AREA AND DATA

The Yellow River is the second longest river in China and originates in the northern part of the Bayankala Mountains in Qinghai Province. The length of the Yellow River is 5,464 km with a drainage area of 752,440 km^2, it passes through nine provinces and autonomous regions and reaches the Bohai Sea in Shangdong Province. The map of the Yellow River basin is shown in Figure 1. The Yellow River has 58 billion m^3 of runoff (during the period of 1919–1975), 59% of which comes from the upper reaches, 33% from the middle reaches, the Loess Plateau. The annual precipitation ranges from 300 mm in the northwest to 700 mm in the southwest (Wang *et al.* 2008; Li *et al.* 2009). About 100 million people reside within the catchment, and the catchment consists of 1,200 million ha of farmland, of which nearly half is irrigated by the Yellow River (Yang *et al.* 2004).

In this study, the observed annual rainfall series from 1956 to 2000 in the sub-water resources region of upper Longyangxia (Subregion 1, Figure 1) and the sub-water resources region between Longyangxia and Lanzhou (Subregion 2, Figure 1) are chosen (Figure 2). Accordingly, the natural annual runoff series from 1956 to 2000 in the Lanzhou station are chosen (Figure 3).

METHODOLOGY

The ensemble empirical mode decomposition (EEMD)

The EMD technique was originally proposed for the study of ocean waves and was a generally non-linear, non-stationary data processing technique (Huang *et al.* 1998). It assumes that original data may have many different coexisting modes of oscillations at the same time owing to its complexity and can be represented by intrinsic mode functions (IMFs) components and a mean value or trend item after smooth treatment. If IMF meets the following two conditions – first, the functions have the same numbers of extrema and zero-crossings or differ at the most by one; second, the functions are symmetric with respect to local zero mean – the EMD can extract these intrinsic modes with different time-scale from the original time series, based on the local characteristic scale of data itself. The

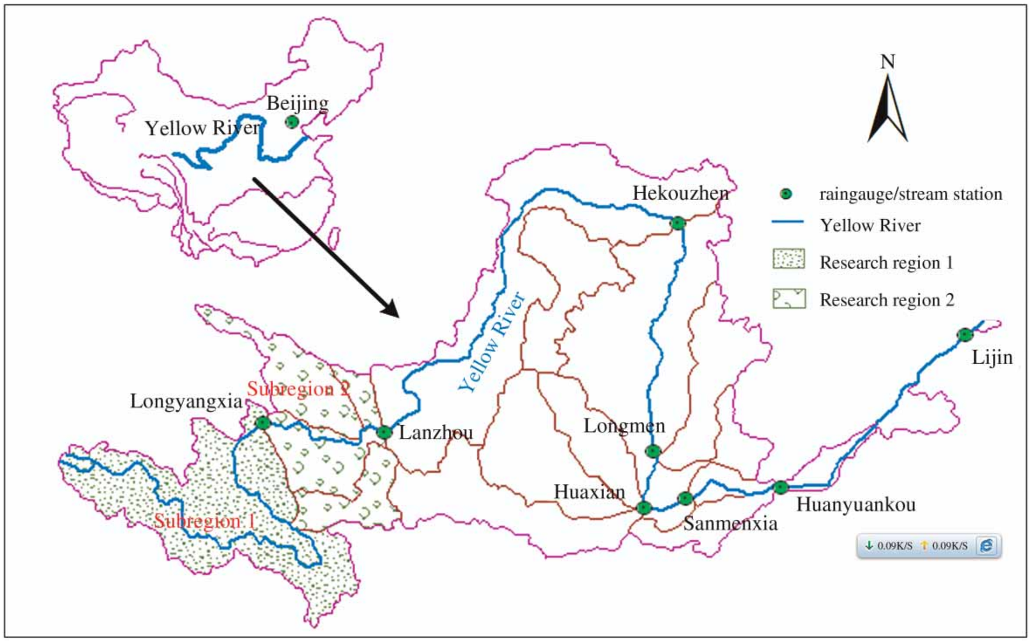

Figure 1 | Location map of the study area.

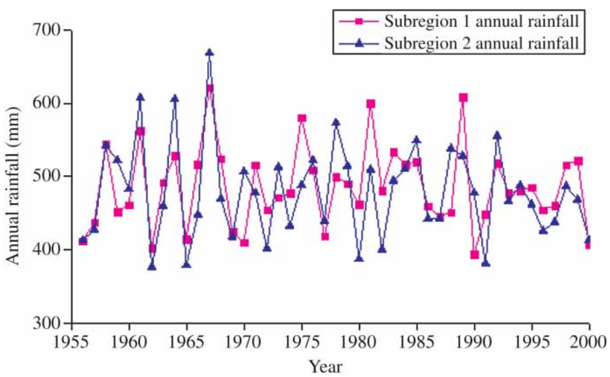

Figure 2 | The annual rainfall series in Subregion 1 and Subregion 2.

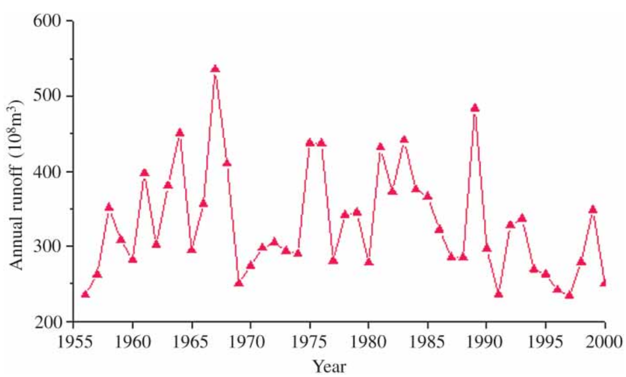

Figure 3 | The annual runoff series in Lanzhou station.

major advantage of EMD is that the basis functions can be derived directly from the data itself depending on the data-driven mechanism which does not require *a priori* knowledge unlike wavelet and Fourier transform.

The EMD algorithm extracts the IMF modes from a given time series through a shifting process (Huang *et al.* 1998). It can be briefly described as follows. (1) Identify all local maxima and minima points for a given time series $x(t)$. (2) Connect all local maxima points to form an upper envelope $e_{\max}(t)$ and all minima points to form a lower envelope $e_{\min}(t)$ with spline interpolation. (3) Calculate the mean $m(t)$ between two envelopes using

$$m(t) = (e_{\max}(t) + e_{\min}(t))/2 \tag{1}$$

(4) Extract the mean from the time series and calculate the difference of $x(t)$ and $m(t)$ as $h(t)$

$$h(t) = x(t) - m(t) \tag{2}$$

(5) If $h(t)$ meets the two conditions of IMF according to stopping criterion, $h(t)$ is denoted as the first IMF [written as $c_1(t)$ and 1 is its index], if $h(t)$ is not an IMF, $x(t)$ is replaced with $h(t)$ and iterate steps 1–4 until $h(t)$ meets

the two conditions of IMF. (6) The residue $r_1(t) = x(t) - c_1(t)$ is then treated as new data subjected to the same sifting process as described above for the next IMF from $r_1(t)$. Finally, the whole decomposition is completed with a finite number of IMFs until the residual satisfies some stopping criteria.

The stopping criterion presented by Huang *et al.* (2003) for extracting an IMF is: iterating predefined times after the residue satisfies the restriction that the number of zero-crossings and extrema do not differ by more than 1. The shifting procedure can be stopped when the residue $r(t)$ becomes a monotonic function or at most has one local extreme point from which no more IMF can be extracted. Having determined successively the different IMFs $c_1(t)$, $c_2(t)$,…, $c_n(t)$ and $r_n(t)$ and the original time series can be rewritten as the sum of some IMFs and a residue:

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t) \tag{3}$$

where n is the number of IMFs, $r_n(t)$ denotes the final residue and $c_i(t)$ are nearly orthogonal to each other, and all have zero means.

For more information regarding the EMD algorithm and stopping criteria of shifting procedure see Huang *et al.* (1998, 2003).

The EMD has been proved quite versatile in a broad range of applications for extracting signals from data generated in noisy non-linear and non-stationary processes (Huang & Shen 2005). However, one major drawback of the original EMD is the frequent appearance of mode mixing, which is defined as a single IMF either consisting of signals of widely disparate scales, or a signal of a similar scale residing in different IMF components (Wu & Huang 2009). To overcome the drawbacks, an ensemble EMD (EEMD) has been developed by Wu & Huang (2009), which defines the true IMF components as the mean of an ensemble of trials, each consisting of the signal plus a white noise of finite amplitude. A complete introduction of EEMD can be found in Wu & Huang (2009). The proposed EEMD is briefly described as follows. (1) Add a white noise series to the targeted data. (2) Decompose the data with added white noise into IMFs. (3) Repeat steps 1 and 2 again and again, until the residue $r(t)$ becomes a monotonic function or at most has one local extreme point from which no more IMF can be extracted, but with a different white noise series each time. (4) Obtain the (ensemble) means of corresponding IMFs of the decompositions as the final result.

The support vector machines (SVM)

SVM is a statistical regression method developed by Vapnik (1995). It uses a linear model to separate sample data through some non-linear mapping from the input vectors into the high-dimensional feature space according to the structural risk minimization (SRM) principle. The most important principle of SVM is the application of minimizing an upper bound to the generalization error instead of minimizing the training error. Based on this, SVM can achieve an optimum networks structure. The detailed description of the theory of SVM has been provided by many books and papers (Vapnik 1998; Bray & Han 2004; Steinwart & Chrismann 2008; Wang *et al.* 2009; Guo *et al.* 2011), and hence only a brief description of SVM is given here. The basic idea of SVM for regression is to map non-linearly the original data x into a high-dimensional feature space and then to perform a linear regression in the feature space.

Given a set of training data $\{(x_i, d_i)\}_i^N$ (x_i is the input vector, d_i is the actual value and N is the total number of data patterns), the general SVM regression function is:

$$y = f(x) = w\varphi(x_i) + b \tag{4}$$

where $\varphi(x_i)$ represents the high dimensional feature spaces, which is non-linearly mapped from the input space x, and w and b are the weight vector and bias term, respectively (Vapnik 1998). w and b can be estimated by minimizing the error function (Equation (5)) and introducing the positive slack variables ξ and ξ^*.

$$\text{Minimize: } \frac{1}{2}||w||^2 + C\left(\sum_{i}^{N}(\xi_i + \xi_i^*)\right) \tag{5}$$

$$\text{subject to } \begin{cases} w_i\varphi(x_i) + b_i - d_i \le \varepsilon + \xi_i^*, \ i = 1, 2, \ldots, N \\ d_i - w_i\varphi(x_i) - b_i \le \varepsilon + \xi_i, \ i = 1, 2, \ldots, N \\ \xi_i, \xi_i^*, i = 1, 2, 3, \ldots, N \end{cases}$$

where $(1/2)||w||^2$ is the weights vector norm and C is referred to as the regularized constant determining the tradeoff between the empirical error and the regularized term. Increasing the value of C will result in an increasing relative importance of the empirical risk with respect to the regularization term. ε is called the tube size and is equivalent to the approximation accuracy placed on the training data points. Both C and ε are user-determined parameters.

By introducing Lagrange multipliers α_i and α_i^*, the above mentioned optimization problem is transformed into the dual quadratic optimization problem. After the quadratic optimization problem with inequality constraints is solved, the parameter vector w in Equation (4) can be obtained:

$$\omega^* = \sum_{i=1}^{N} (\alpha_i - \alpha_i^i)\varphi(x_i) \tag{6}$$

Therefore, the SVR regression function is obtained as Equation (7):

$$f(x, \alpha, \alpha^*) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*)K(x, x_i) + b \tag{7}$$

Here, $K(x, x_i)$ is called the Kernel function. The value of the Kernel is inner product of the two vectors x and x_i in the feature space $\varphi(x)$ and $\varphi(x_i)$, so $K(x, x_i) = \varphi(x) \times \varphi(x_i)$, and a function that satisfies Mercer's condition (Vapnik 1998) can be used as the Kernel Function. In general, there are several types of kernel function, namely linear, polynomial and radial basis function (RBF). The most used kernel function is the RBF, as follows:

$$K(x, x_i) = \exp\left(-||x - x_i||^2/2\sigma^2\right) \tag{8}$$

The RBF kernel has been reported as the best choice over other kernel functions (Dibike et al. 2001; Noori et al. 2009). It is not only capable to map non-linearly the training data into an infinite dimensional space, but also easier to implement. Therefore, the RBF kernel function is employed to deal with non-linear relationship problems in this study.

The selection of the parameters C, ε and σ has a great influence on the forecasting accuracy of a SVR model. For example, the parameter C determines the tradeoff cost between minimizing the training error and minimizing model complexity. If C is too large (infinity), then the objective is to minimize the empirical risk only. Parameter ε controls the width of the e-insensitive loss function. Large ε-values result in a flatter regression estimated function. Parameter σ controls the RBF width, which reflects the distribution range of x-values of training data and it is also user-determined parameters. Thus, the determination of all three parameters selection technique is an important issue. Some practical analytical approaches were described by Cherkassky & Ma (2004) to the selection of C and ε based on *a priori* knowledge and experience, cross-validation, and asymptotical optimization. However, there is no structural method for efficiently and simultaneously confirming those parameters of SVR model. Therefore, a PSO is used to optimize the parameters of SVM model.

The particle swarm optimization (PSO)

The PSO developed by Kenndey & Eberhart (1995) is an evolutionary computation technique, which is inspired in the emerging properties of collective behavior of some animals, such as bird flocking, bee swarming, and fish schooling. Due to its easy implementation and inexpensive computation, its simplicity in coding and consistency in performance, the PSO has proved to be an effective and competitive algorithm to solve variable optimization problems (Chau 2006; Schwaab et al. 2008; Zhu et al. 2011).

In the PSO algorithm, each individual of the population is termed the particle and members of the population are termed the swarm. In order to solve an optimization problem, the initial position and velocity of each particle are generated randomly. During the solution process, each particle has an adaptable position and velocity according to which it moves in the search space. Moreover, each particle can remember the best position of the search space it has ever visited. A brief statement of the PSO algorithm can be given as follows.

Let the position and velocity of the ith particle in the m-dimensional search space be represented as $X_i = (x_{i1}, x_{i2}, \ldots x_{im})$, and $V_i = (v_{i1}, v_{i2}, \ldots, v_{im})$, respectively. The best previous position of the ith particle is recorded and represented as $P_i = (p_{i1}, p_{i2}, \ldots, p_{im})$. g is defined as the index of the best particle among all the particles in the group. The modified velocity and position of each particle can be calculated according to the following two equations:

$$v_{im}^{(k+1)} = wv_{im}^{(k)} + c_1 r_1 (p_{im} - x_{im}^k) + c_2 r_2 (p_{gm} - x_{im}^k) \tag{9}$$

$$x_{im}^{(k+1)} = x_{im}^k + v_{im}^{(k+1)}, \quad i = 1, 2, \ldots, n \tag{10}$$

where k is the iteration index, w is inertia weight, c_1 and c_2 are the acceleration coefficients which are used to determine how much the particle's personal best and the global best influence its movement, r_1 and r_2 are uniform random numbers between 0 and 1, v_{im}^k is the current velocity of the particle at iteration k, x_{id}^k is the current position of the particle at iteration k, $v_{im}^{(k+1)}$ is the modified velocity, $x_{im}^{(k+1)}$ is the position of the particle at iteration $k+1$, n is the number of particles in a population.

According to past experiences, the acceleration constants c_1 and c_2 were often set to be 2.0. A larger value of w leads to global exploration, whereas smaller values result in a fine search within the solution space. Therefore, suitable selection of inertia weight w provides a balance between global and local explorations. In general, the inertia weight w is set according to the following equation:

$$w = w_{\max} - \frac{w_{\max} - w_{\min}}{\text{iter}_{\max}} \times \text{iter} \tag{11}$$

where $\text{iter}_{\max}$ is the maximum number of iterations and iter is the current number of iterations.

Shi & Eberhart (1998) found a significant improvement in the performance of PSO with the linearly decreasing inertia weight over the generations, given as the following:

$$w = (w_{\max} - w_{\min}) \times \left(\frac{\text{iter}_{\max} - \text{iter}}{\text{iter}_{\max}}\right) + w_{\min} \tag{12}$$

where $w_{\max}$ and $w_{\min}$ are the higher and lower inertia weight values and the values of w will decrease from $w_{\max}$ to $w_{\min}$. iter is the current iteration and $\text{iter}_{\max}$ is the maximum number of iterations.

The artificial neural network

The ANN model is a flexible mathematical structure patterned after the biological nervous system (Yoon et al. 2011). Many studies in the context of rainfall-runoff process using ANNs have been carried out (Hsu et al. 1995; Sajikumar & Thandaveswara 1999; Tokar & Johnson 1999; Lin & Chen 2004; Srinivasulu & Jain 2006; Wu & Chau 2011). Hence, a three-layer feed-forward ANN model trained with scaled conjugate gradient algorithm (Moller 1993) is used for rainfall-runoff modeling. The basic structure of a feed-forward ANN used can be seen in Figure 4. It consists of a number of neurons arranged in different layers: the input layer, where the data are introduced to the network; the hidden layer or layers, where data are processed; and the output layer, where the results of given input are produced. In the modeling process, all the data series were normalized using the minimum ($X_{\min}$) and maximum ($X_{\max}$) values as described in Equation (13), so that the variables value set ranged from 0 to 1. The tan-sigmoid transfer function is adopted in determining the neurons of the hidden layer whilst the LTF is used in determining the neurons of the output layer. The training epoch is set to 1,000.

$$\text{normalized } x = \frac{X - X_{\min}}{X_{\min}}$$

$$\text{normalized } x = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{13}$$

The ordinary least-squares (OLS) regression

OLS regression (Hutcheson & Sofroniou 1999) is a generalized linear modeling technique that is widely used to predict

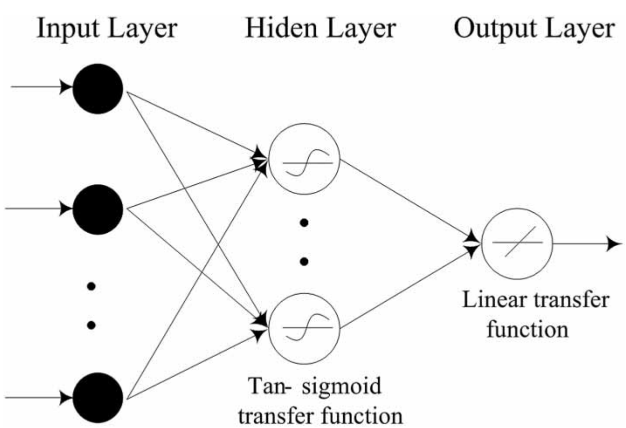

Figure 4 | Architecture of three layers feed-forward back-propagation ANN.

values of a continuous response variables using a single or multiple explanatory variables. In OLS regression, the relationship between a continuous response variables (Y) and a continuous explanatory variable (X) may be appropriately represented mathematically using the straight line equation $Y = \alpha + \beta X$. The OLS regression model can be extended to include multiple explanatory variables by simply adding variables to the equation (Thorpe & Holt 2007). The form of the model is the same as above with a single response variable (Y), but the equation predicting Y by multiple explanatory variables (X_1, X_2) is obtained as $Y = \alpha + \beta_1 X_1 + \beta_2 X_2$. Moreover, the parameters (α and β) can be calculated by determining the equation that minimizes the sum of the squared distances between the observed data and values predicted by the equation. In this study, the equation of runoff predicting Y by two annual rainfall series variables is obtained as:

$$Y = -210.8743 + 0.9358X_1 + 0.1798X_2 \tag{14}$$

MODEL CONSTRUCTION

Input data-preprocessing with EEMD

The EEMD technique is used to preprocess the annual rainfall series from 1956 to 2000 in Subregion 1 and Subregion 2. Before using the EEMD, the two parameters of the number of the ensemble and the amplitude of white noise need to be set. Wu & Huang (2009) pointed out that the effect of the added white noise should decrease following the well-established statistical rule:

$$e_n = \frac{\varepsilon}{\sqrt{N}} \tag{15}$$

where N is the number of ensemble members, ε denotes the amplitude of the added noise, and e_n means the standard deviation of error, which is defined as the difference between the input signal and the corresponding IMFs (Wu & Huang 2009). In the application, if the added noise amplitude is too small, then it may not cause the change of extrema that the EMD depends on. Instead, if the added noise is too large, it would result in redundant IMF components. In this paper, an ensemble member of the EEMD method is set as 100 and the added white noise in each ensemble member has a standard deviation of 0.2. These parameters are similar to those employed by Wu & Huang (2009) where the reader can find a detailed analysis of noise and ensemble averaging effects on decomposition, hence these are not repeated here.

Using the EEMD technique previously described, results are shown in Figures 5–9. There are four IMFs components (Figures 5–8) and a trend term, which is the residue (Res) component (Figure 9). These results are used as input data to improve the prediction performance.

From Figures 5–8, we can also see that the IMFs present changing frequencies, amplitudes and wavelengths. IMF1 is the maximum amplitude, highest frequency and shortest wavelength. The following IMF components decrease in

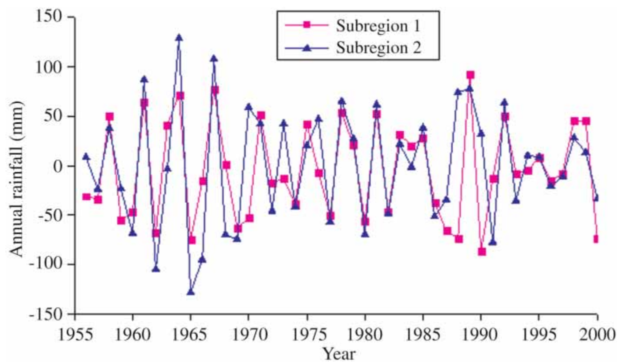

Figure 5 | IMF1 component of annual rainfall series.

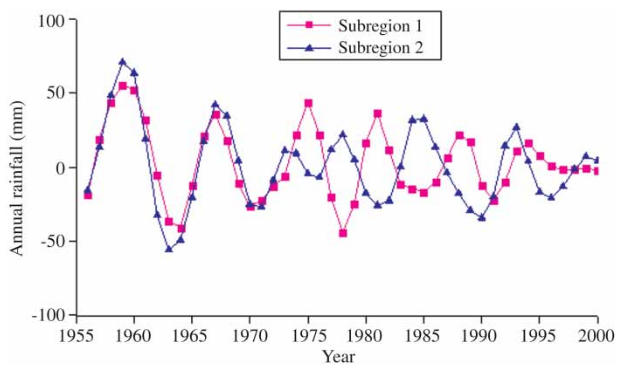

Figure 6 | IMF2 component of annual rainfall series.

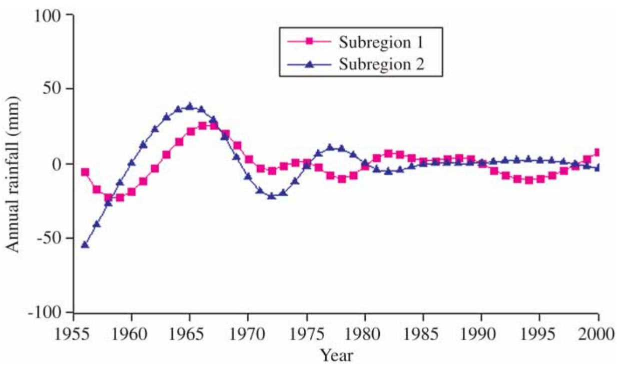

Figure 7 | IMF3 component of annual rainfall series.

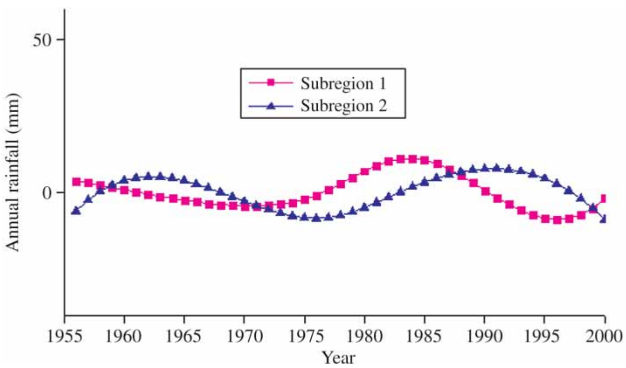

Figure 8 | IMF4 component of annual rainfall series.

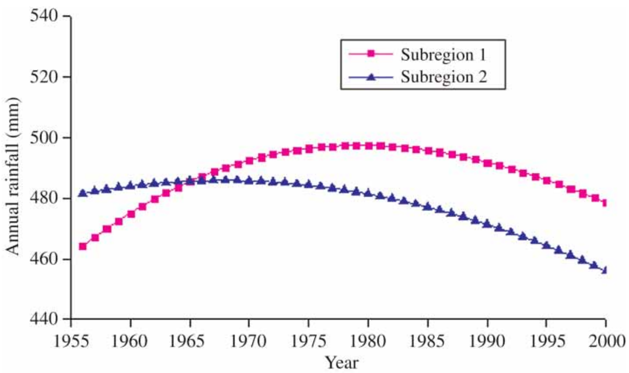

Figure 9 | Res component of annual rainfall series.

the amplitude and frequency, and increase in the wavelength. The last residue is a mode slowly varying around the long term average. The IMF1 component of Subregion 1 and the IMF2 component of Subregion 2 in Figures 5 and 6 show that there is obvious periodic variability within 3–4 years and 7–8 years, respectively. According to multiple timescale analysis of sea surface temperature (SST) data in the last 100 years (Sun & Lin 2006), there is also obvious periodic variability within 3–4 years and 7–8 years. Hence, the high frequency components of annual rainfall series are consistent with SST and demonstrate that the short-term variation of annual rainfall in the study area may be affected by SST. The amplitude of IMF3 shows a steady decreasing tendency from 1970 to 2000 and demonstrates that the amplitude range of annual rainfall becomes smaller after 1970. IMF4 in Figure 8 has an average period of 22 years. Res component shows the overall trend of annual rainfall data in Subregion 1 and Subregion 2 in Figure 9. The trend of annual rainfall in Subregion 1 shows an increasing tendency from 1956 to 1980, and then shows decreasing tendency from 1980 onwards. The trend of annual rainfall in Subregion 2 shows a decreasing tendency.

Optimizing the SVM parameters with PSO

Here, the particle is composed of the parameters C, ε and σ. Figure 10 presents the process of optimizing the SVM parameters with PSO algorithm, which is described as

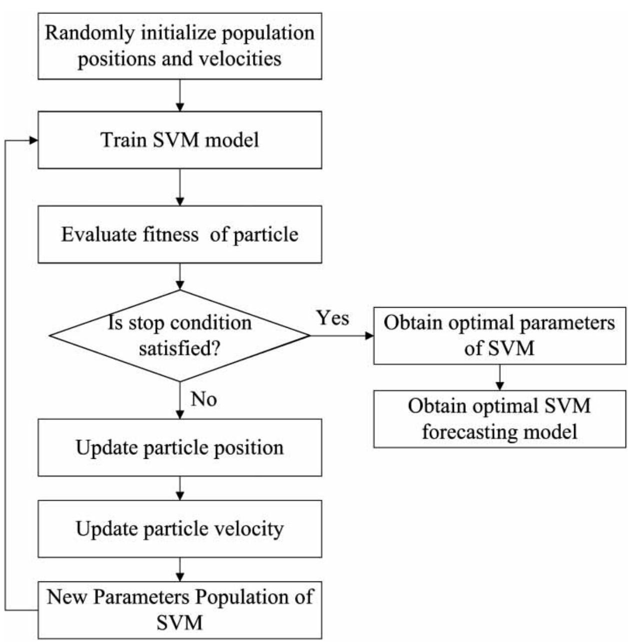

Figure 10 | The flowchart of optimizing the SVM parameters with PSO.

follows. (1) PSO is initialized with a population of random positions and velocities in the given value range of SVM parameters. (2) SVM model is trained with parameters C, ε and σ included in current particle. (3) Evaluate fitness of particle. In order to evaluate fitness of particle, the root mean squared error (RMSE) is used as the fitness function in this study. It is one of the commonly used error index statistics (Lin *et al.* 2006; Wang *et al.* 2009). (4) If the stopping criterion is satisfied, the optimal parameters of SVM are obtained. Otherwise, the new parameters population of SVM can be obtained by updating particle position and velocity, hence go to (2) until the stopping criterion is satisfied.

In this paper, the SVM model is constructed for forecasting annual runoff time series. The 10 components of annual rainfall which are obtained by using EEMD technique are used as input data and the corresponding annual runoff data are used as output data. The dataset from 1956 to 1995 is used for training the model whilst that from 1996 to 2000 is used for prediction. In order to obtain the optimal parameters of the SVM mode, the PSO algorithm previously introduced is employed. Concerning its implementation, we need to set five parameters of the PSO algorithm: inertia weight $w_{\max}$ and $w_{\min}$, the acceleration constants c_1 and c_2, and the maximum number of iterations $\text{iter}_{\max}$. The inertia weight provides a balance between global and local explorations, and is used to control the convergence behavior of the PSO. As originally developed, the inertia weight often decreases linearly from about 0.9 to 0.4 during a run (Sadati *et al.* 2009). The acceleration coefficients c_1 and c_2 control how far a particle will move in a single iteration. Low values allow particles to roam far from target regions before being tugged back, while high values result in abrupt movement towards, or past, target regions (Kenndey & Eberhart 1995). Typically, these are both set to a value of 2.0, although assigning different values to c_1 and c_2 sometimes leads to improved performance (Poli *et al.* 2007). Hence, the parameters of the PSO are set as follows: inertia weight $w_{\max} = 0.9$, $w_{\min} = 0.4$, the acceleration constants $c_1 = c_2 = 2.0$, maximum iterations $\text{iter}_{\max} = 500$. SVM optimal parameters $(C, \varepsilon, \sigma) = (4.9851, 0.0313, 0.9195)$ were obtained for rainfall-runoff forecasting using SVM based on EEMD by PSO.

In order to measure the forecasting capability of the proposed PSO–SVM based on EEMD decomposed annual rainfall series (I), a PSO–SVM model with the original annual rainfall series (II) and optimal parameters $(C, \varepsilon, \sigma) = (1.6633, 0.0313, 0.0122)$ obtained by PSO, a typical three-layer feed-forward ANNs based on EEMD decomposed annual rainfall series (III), a typical three-layer feed-forward ANNs with the original annual rainfall series (IV) and an OLS regression (V) are used as the benchmark models.

PERFORMANCE EVALUATION

To measure the forecasting performance of the models, four main criteria are used for evaluation of level prediction and directional forecasting, respectively. First, the RMSE is selected as the performance criterion of level prediction. As one of the commonly used error index statistics, the RMSE is defined as

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Q_f(i) - Q_0(i))^2} \tag{16}$$

Second, the average absolute relative error (AARE) is selected as the accuracy criterion of level prediction. It is an unbiased statistic for measuring the predictive capability of a model. The AARE is defined as

$$\text{AARE}(\%) = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Q_f(i) - Q_0(i)}{Q_f(i)}\right| \times 100$$

$$\text{AARE}(\%) = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Q_f(i) - Q_0(i)}{Q_0(i)}\right| \times 100 \tag{17}$$

Third, the coefficient of correlation (R) is selected as the degree of collinearity criterion of level prediction. It is given as

$$R = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(Q_0(i) - \overline{Q_0}\right)\left(Q_f(i) - \overline{Q_f}\right)}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Q_0(i) - \overline{Q_0}\right)^2} * \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Q_f(i) - \overline{Q_f}\right)^2}} \tag{18}$$

Finally, the Nash–Sutcliffe efficiency (NSE) is a very popular index to assess the predictive power of hydrological models. It is defined as (Nash & Sutcliffe 1970)

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n}(Q_0(i) - Q_f(i))^2}{\sum_{i=1}^{n}\left(Q_0(i) - \overline{Q_0}\right)^2} \tag{19}$$

where $Q_0(i)$ and $Q_f(i)$ are the observed and forecasted discharge, respectively, and $\overline{Q_0}$, $\overline{Q_f}$ denote their means, and n is the number data points considered.

APPLICATION AND COMPARISON

For the same basis of comparison, the same training and testing sets are used for all the above five models developed, and the prediction results and relative errors obtained from those five models for the rainfall-runoff forecasting in Lanzhou station are presented in Table 1. The four quantitative standard statistical performance evaluation measures RMSE, AARE, R and NSE are employed to evaluate the performances of the five models developed, and the statistical results of different models are summarized in Table 2. From Table 2, it can be observed the value with PSO–SVM and ANN based on EEMD decomposed annual rainfall series were able to produce a good and close forecast, as compared with those of the PSO–SVM and ANN with the original annual rainfall series. The PSO–SVM–EEMD model improved the PSO–SVM model, ANN–EEMD model, ANN model forecasting, OLS regression by about 26.27, 53.73, 65.99 and 64.40%, respectively, and a 28.94, 53.72, 67.7, and 65.38% reduction in RMSE and AARE values, respectively. Improvements in the forecasting results regarding the R value were approximately 5.88, 34.33, 8.43, and 18.89%, respectively. The PSO–SVM–EEMD model obtained the best value of NSE increases by 20.99 and 87.64% comparing with the PSO–SVM model and ANN–EEMD model forecasting, respectively, and the ANN model and OLS regression obtained the negative value of NSE which indicate that the observed mean is a better predictor than them. In addition, the obtained value of RMSE and AARE in the ANN–EEMD model decrease by 26.49 and 27.95%, respectively, comparing with their counterparts in the ANN model. Figure 11 illustrates the rainfall-runoff forecasting results using different models. It can be seen from Figure 11 that the PSO–SVM–EEMD model can match runoff better than PSO–SVM, ANN–EEMD and ANN models. The results of this analysis indicate that the PSO–SVM–EEMD model is able to obtain the best result in terms of different evaluation measures. This illustrates that the EEMD model is suitable for non-linear and non-stationary hydrologic data analysis, the idea of 'decomposition and ensemble' is feasible, and the hybrid PSO–SVM–EEMD model can overcome the drawbacks of individual models to generate a synergetic effect in forecasting. Thus the annual rainfall data decomposed using the EEMD technique as input data of models can improve the prediction performance.

Table 1 | Comparison of forecasting value of relative error of five rainfall-runoff models during testing period

		Forecasted value (10^8 m^3)					Relative error (%)				
Year	Observed value (10^8 m^3)	I	II	III	IV	V	I	II	III	IV	V
1996	242.0	234.7	234.0	251.5	283.5	290.89	– 3.02	– 3.31	3.93	17.17	20.20
1997	234.0	252.9	261.9	251.8	289.4	298.62	8.08	11.91	7.61	23.67	27.62
1998	279.2	308.7	236.1	313.4	372.1	359.39	10.57	– 15.45	12.25	33.27	28.72
1999	348.4	328.7	338.0	317.2	378.2	361.95	– 5.65	– 2.98	– 8.96	8.55	3.89
2000	250.6	247.4	234.0	177.8	265.4	245.10	– 1.28	– 6.62	– 29.05	5.90	– 2.19

I – PSO–SVM based on EEMD decomposed annual rainfall series.
II – PSO–SVM model with the original annual rainfall series.
III – typical three-layer feed-forward artificial neural networks based on EEMD decomposed annual rainfall series.
IV – typical three-layer feed-forward artificial neural networks with the original annual rainfall series.
V – OLS regression.

Table 2 | Forecasting performance indices of models

Model	Forecasting performance RMSE	AARE (%)	*R*	NSE
I	18.33	5.72	0.90	0.81
II	24.85	8.05	0.85	0.64
III	39.62	12.36	0.67	0.10
IV	53.90	17.71	0.83	– 0.67
V	51.49	16.52	0.73	– 0.52

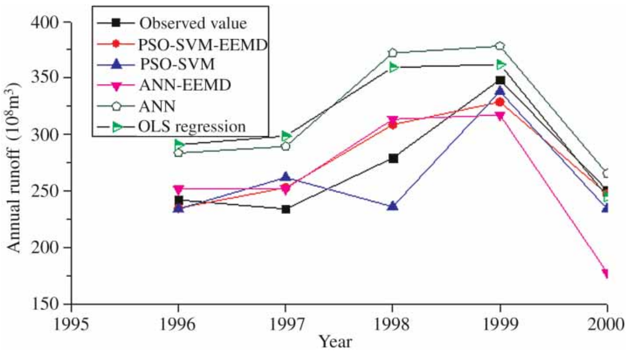

Figure 11 | Observed and forecasted runoff during testing period by five models.

CONCLUSIONS

The study reported in this paper investigates the use of a new adaptive data analysis methodology, EEMD, for decomposing annual rainfall series in rainfall-runoff model based on SVM. The parameters of SVM are determined by PSO, which is not needed to consider the analytic property of the generalization performance measure and can avoid the occurrence of over-fitting or under-fitting of the SVM model due to improper determination of these parameters. The proposed method and models were tested using real datasets from a large size catchment of the Yellow River in China. Using EEMD technique, the original annual rainfall series can be decomposed into four independent IMFs and one residue, which were used as input variables for modeling rainfall-runoff process. A typical three-layer feed-forward ANN model and OLS regression are also employed for illustrating the forecasting capability of the proposed model. Four standard statistical performance evaluation measures are adopted to evaluate the performances of various models developed. The empirical results indicate that the PSO–SVM model based on the EEMD approach can enhance significantly rainfall-runoff forecasting at the studied station in the Yellow River.

ACKNOWLEDGEMENTS

This research was supported by the Key Project of Science Technique Research of Henan Educational Committee (No. 12A570005), the foundation for University Backbone Teacher of Henan Province (2012GGJS-099) and Central Research Grant of Hong Kong Polytechnic University (4-ZZAD). We gratefully acknowledge the critical comments and corrections of the anonymous reviewers.

REFERENCES

Abrahart, R. J. 2003 Neural network rainfall–runoff forecasting based on continuous resampling. *Journal of Hydroinformatics* **5**, 51–61.
Anmala, J., Zhang, B. & Govindaraju, R. S. 2000 Comparison of ANNs and empirical approaches for predicting watershed runoff. *Journal of Water Resources Planning and Management* **126**, 156–166.
ASCE Task Committee on Application of the Artificial Neural Networks in Hydrology 2000a Artificial neural networks in hydrology I: preliminary concepts. *Journal of Hydrologic Engineering* **5**, 115–123.
ASCE Task Committee on Application of the Artificial Neural Networks in Hydrology 2000b Artificial neural networks in hydrology II: hydrologic applications. *Journal of Hydrologic Engineering* **5**, 124–137.
Asefa, T., Kemblowski, M., McKee, M. & Khalil, A. 2006 Multi-time scale stream flow predictions: the support vector machines approach. *Journal of Hydrology* **318**, 7–16.
Behzad, M., Asghari, K., Eazi, M. & Palhang, M. 2009 Generalization performance of support vector machines and neural networks in runoff modeling. *Expert Systems with Applications* **36**, 7624–7629.
Braddock, R. D., Kremmer, M. L. & Sanzogni, L. 1998 Feed-forward artificial neural network model for forecasting rainfall run-off. *Environmetrics* **9**, 419–432.
Bray, M. & Han, D. 2004 Identification of support vector machines for runoff modeling. *Journal of Hydroinformatics* **6**, 265–281.
Chau, K. W. 2006 Particle swarm optimization training algorithm for ANNs in stage prediction of Shing Mun River. *Journal of Hydrology* **329**, 363–367.

Cherkassky, V. & Ma, Y. 2004 Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Networks* **17**, 113–126.

Coughlin, K. T. & Tung, K. K. 2004 11-Year solar cycle in the stratosphere extracted by the empirical mode decomposition method. *Advances in Space Research* **34**, 323–329.

Dawson, C. W. & Wilby, R. 1998 An artificial neural network approach to rainfall–runoff modeling. *Hydrological Sciences Journal* **43**, 47–66.

Dawson, C. W. & Wilby, R. L. 2001 Hydrological modelling using artificial neural networks. *Progress in Physical Geography* **25**, 80–108.

Dibike, Y. B., Velickov, S., Solomatine, D. P. & Abbott, M. B. 2001 Model induction with support vector machines: introduction and applications. *Journal of Computing in Civil Engineering* **15**, 208–216.

Guo, Z. H., Zhao, W. G., Lu, H. Y. & Wang, J. Z. 2012 Multi-step forecasting for wind speed using a modified EMD based artificial neural network model. *Renewable Energy* **37**, 241–249.

Guo, J., Zhou, J., Qin, H., Zou, Q. & Li, Q. 2011 Monthly streamflow forecasting based on improved support vector machine model. *Expert Systems with Applications* **38**, 13073–13081.

Halff, A. H., Halff, H. M. & Azmoodeh, M. 1993 Predicting runoff from rainfall using neural networks. In *Proc. Eng. Hydrol.*, ASCE, New York, pp. 768–775.

Hsu, K. L., Gupta, H. V. & Sorooshian, S. 1995 Artificial neural network modeling of the rainfall–runoff process. *Water Resources Research* **31**, 2517–2530.

Huang, N. E. & Shen, S. S. P. 2005 *Hilbert–Huang Transform and Its Applications*. World Scientific, Singapore.

Huang, Y., Schmitt, F. G., Lu, Z. & Liu, Y. 2009 Analysis of daily river flow fluctuations using empirical mode decomposition and arbitrary order Hilbert spectral analysis. *Journal of Hydrology* **373**, 103–111.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. L. C., Shih, H. H., Zheng, Q. N., Yen, N. C., Tung, C. C. & Liu, H. H. 1998 The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London Series A-Mathematical Physical and Engineering Sciences* **454**, 903–995.

Huang, N. E., Wu, M. L., Long, S. R., Shen, S. S. P., Qu, W., Gloersen, P. & Fan, K. L. 2003 A confidence limit for the empirical mode decomposition and Hilbert spectral analysis. *Proceedings of the Royal Society of London Series A-Mathematical Physical and Engineering Sciences* **459**, 2317–2345.

Hutcheson, G. D. & Sofroniou, N. 1999 *The Multivariate Social Scientist: Introductory Statistics Using Generalized Linear Models*. Sage Publications Ltd, London.

Jeong, D. & Kim, Y. O. 2005 Rainfall–runoff models using artificial neural networks for ensemble streamflow prediction. *Hydrological Processes* **19**, 3819–3835.

Karagiannis, A. & Constantinou, P. 2011 Noise-assisted data processing with empirical mode decomposition in biomedical signals. *IEEE Transactions on Information Technology in Biomedicine* **15**, 11–18.

Karunanithi, N., Grenney, W. J., Whitley, D. & Bovee, K. 1994 Neural networks for river flow prediction. *Journal of Computing in Civil Engineering, ASCE* **8**, 201–220.

Kenndey, J. & Eberhart, R. 1995 Particle swarm optimization. In *IEEE International Conference on Neural Networks Proceedings*, Piscataway, NJ, pp. 1942–1948.

Li, C. H., Yang, Z. F., Huang, G. H. & Li, Y. P. 2009 Identification of relationship between sunspots and natural runoff in the Yellow River based on discrete wavelet analysis. *Expert Systems with Applications* **36**, 3309–3318.

Lin, G. F. & Chen, L. H. 2004 A non-linear rainfall-runoff model using radial basis function network. *Journal of Hydrology* **289**, 1–8.

Lin, J. Y., Cheng, C. T. & Chau, K. W. 2006 Using support vector machines for long-term discharge prediction. *Hydrological Sciences Journal* **51**, 599–612.

Lin, G. F. & Wu, M. C. 2011 An RBF network with a two-step learning algorithm for developing a reservoir inflow forecasting model. *Journal of Hydrology* **405**, 439–450.

Liong, S. Y. & Sivapragasam, C. 2002 Flood stage forecasting with support vector machines. *Journal of the American Water Resources Association* **38**, 173–186.

Liong, S. Y., Gautam, R. R. & Khu, S. T. 2002 Genetic Programming: a new paradigm in rainfall runoff modeling. *Journal of the American Water Resources Association* **38**, 705–718.

Lohani, A. K., Goel, N. K. & Bhatia, K. K. S. 2011 Comparative study of neural network, fuzzy logic and linear transfer function techniques in daily rainfall-runoff modelling under different input domains. *Hydrological Processes* **25**, 175–193.

Maier, H. & Dandy, G. C. 2000 Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications. *Environmental Modeling and Software* **15**, 101–124.

Minns, A. W. & Hall, M. J. 1996 Artificial neural networks as rainfall–runoff models. *Hydrological Sciences Journal* **41**, 399–417.

Moller, M. F. 1993 A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks* **6**, 525–533.

Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models part I – a discussion of principles. *Journal of Hydrology* **10**, 282–290.

Noori, R., Abdoli, M. A., Ghasrodashti, A. A. & Jalili Ghazizade, M. 2009 Prediction of municipal solid waste generation with combination of support vector machine and principal component analysis: a case study of Mashhad. *Environmental Progress & Sustainable Energy* **28**, 249–258.

Nourani, V., Kisi, Ö. & Komasi, M. 2011 Two hybrid artificial intelligence approaches for modeling rainfall-runoff process. *Journal of Hydrology* **402**, 41–59.

Poli, R., Kennedy, J. & Blackwell, T. 2007 Particle swarm optimization: an overview. *Swarm Intelligence* **1**, 33–57.

Pramanik, N., Panda, R. K. & Singh, A. 2011 Daily river flow forecasting using wavelet ANN hybrid models. *Journal of Hydroinformatics* **13**, 49–63.
Sadati, N., Amraee, T. & Ranjbar, A. M. 2009 A global particle swarm-based-simulated annealing optimization technique for under-voltage load shedding problem. *Applied Soft Computing* **9**, 652–657.
Sajikumar, N. & Thandaveswara, B. S. 1999 A non-linear rainfall-runoff model using an artificial neural network. *Journal of Hydrology* **216**, 32–55.
Samsudin, R., Saad, P. & Shabri, A. 2011 River flow time series using least squares support vector machines. *Hydrology and Earth System Sciences* **15**, 1835–1852.
Sarangi, A., Madramootoo, C. A., Enright, P., Prasher, S. O. & Patel, R. M. 2005 Performance evaluation of ANN and geomorphology-based models for runoff and sediment yield prediction for a Canadian watershed. *Current Science* **89**, 2022–2033.
Schwaab, M., Biscaia, J. E. C., Monteiro, J. L. & Pinto, J. C. 2008 Nonlinear parameter estimation through particle swarm optimization. *Chemical Engineering Science* **63**, 1542–1552.
Sedki, A., Ouazar, D. & El Mazoudi, E. 2009 Evolving neural network using real coded genetic algorithm for daily rainfall-runoff forecasting. *Expert Systems with Applications* **36**, 4523–4527.
Shamseldin, A. Y. 1997 Application of a neural network technique to rainfall–runoff modelling. *Journal of Hydrology* **199**, 272–294.
Shamseldin, A. Y., O'Connor, K. M. & Liang, G. C. 1997 Methods for combining the outputs of different rainfall-runoff models. *Journal of Hydrology* **197**, 203–229.
Shi, Y. & Eberhart, R. C. 1998 A modified particle swarm optimizer. In *Proceedings of the IEEE congress on Evolutionary Computation*, Piscataway, NJ, pp. 69–73.
Spanos, P. D., Giaralis, A. & Politis, N. P. 2007 Time-frequency representation of earthquake accelerograms and inelastic structural response records using the adaptive chirplet decomposition and empirical mode decomposition. *Soil Dynamics and Earthquake Engineering* **27**, 675–689.
Srinivasulu, S. & Jain, A. 2006 A comparative analysis of training methods for artificial neural network rainfall-runoff models. *Applied Soft Computing* **6**, 295–306.
Steinwart, I. & Chrismann, A. 2008 *Support Vector Machines*. Springer, Berlin.
Sun, X. & Lin, Z. S. 2006 A new technology HHT and its diagnosis for ENSO. *Meteorological Monthly* **32**, 17–22.
Thorpe, R. & Holt, R. 2007 *The Sage Dictionary of Qualitative Management Research*. Sage Publications Ltd, London.
Tokar, A. S. & Johnson, P. A. 1999 Rainfall–runoff modeling using artificial neural networks. *Journal of Hydrologic Engineering* **4**, 232–239.
Vapnik, V. 1995 *The Nature of Statistical Learning Theory*. Springer, New York.
Vapnik, V. 1998 *Statistical Learning Theory*. Wiley, New York.
Wang, W.-C., Chau, K. W., Cheng, C.-T. & Qiu, L. 2009 A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *Journal of Hydrology* **374**, 294–306.
Wang, S. Y., Liu, J-S. & Yang, C. J. 2008 Eco-environmental vulnerability evaluation in the Yellow River Basin, China. *Pedosphere* **18**, 171–182.
Wu, C. L. & Chau, K. W. 2011 Rainfall-runoff modeling using artificial neural network coupled with singular spectrum analysis. *Journal of Hydrology* **399**, 394–409.
Wu, Z. & Huang, N. E. 2009 Ensemble empirical mode decomposition: a noise assisted data analysis method. *Advances in Adaptive Data Analysis* **1**, 1–49.
Wu, J. D. & Tsai, Y. J. 2011 Speaker identification system using empirical mode decomposition and an artificial neural network. *Expert Systems with Applications* **38**, 6112–6117.
Wu, C. L., Chau, K. W. & Li, Y. S. 2008 River stage prediction based on a distributed support vector regression. *Journal of Hydrology* **358**, 96–111.
Yang, D. W., Li, C., Hu, H. P., Lei, Z. D., Yang, S. X., Kusuda, T., Koike, T. & Musiake, K. 2004 Analysis of water resources variability in the Yellow River of China during the last half century using historical data. *Water Resources Research* **40**, W06502.
Yoon, H., Jun, S.-C., Hyun, Y., Bae, G.-O. & Lee, K.-K. 2011 A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *Journal of Hydrology* **396**, 128–138.
Yu, L., Wang, S. Y. & Lai, K. K. 2008 Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. *Energy Economics* **30**, 2623–2635.
Zhang, B. & Govindaraju, R. S. 2000 Prediction of watershed runoff using Bayesian concepts and modular neural networks. *Water Resources Research* **36**, 753–762.
Zhang, B. & Govindaraju, R. S. 2003 Geomorphology-based artificial neural networks (GANNs) for estimation of direct runoff over watersheds. *Journal of Hydrology* **273**, 18–34.
Zhang, X., Lai, K. K. & Wang, S. Y. 2008 A new approach for crude oil price analysis based on Empirical Mode Decomposition. *Energy Economics* **30**, 905–918.
Zhu, H., Wang, Y., Wang, K. & Chen, Y. 2011 Particle Swarm Optimization (PSO) for the constrained portfolio optimization problem. *Expert Systems with Applications* **38**, 10161–10169.

First received 6 April 2012; accepted in revised form 13 March 2013. Available online 24 April 2013